

(19) 日本国特許庁 (J P)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-208950

(P2002-208950A)

(43) 公開日 平成14年7月26日 (2002. 7. 26)

(51) Int. Cl.<sup>7</sup>

H 0 4 L 12/56

識別記号

2 0 0

F I

H 0 4 L 12/56

データベース (参考)

2 0 0 A 5 K 0 3 0

審査請求 未請求 請求項の数 3 O L 外国語出願 (全 20 頁)

(21) 出願番号 特願2001-369972(P2001-369972)

(22) 出願日 平成13年12月4日 (2001. 12. 4)

(31) 優先権主張番号 0 9 / 7 2 8 9 8 8

(32) 優先日 平成12年12月4日 (2000. 12. 4)

(33) 優先権主張国 米国 (U S)

(71) 出願人 501459712

アカラ コーポレーション

Akara Corporation

カナダ ケイ2エル 2エヌ2 オンタリオ

カナダ カティマヴィックロード

150 スイート2000

(72) 発明者 マシュー ロバート ウィリアムズ

カナダ国, ケー2ティー 1シー5, オン

タリオ, カナダ, インスマイル クレセント

7

(74) 代理人 100077517

弁理士 石田 敬 (外4名)

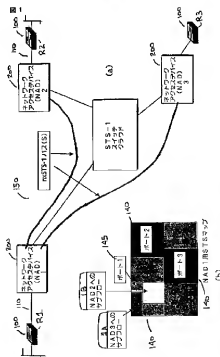
最終頁に続く

(54) 【発明の名称】 N×STS-1 S O N E T上でのパケット転送のための方法および装置

## (57) 【要約】

【課題】 N×STS-1 S O N E Tネットワークに接続されたソースおよび宛先 I/O デバイス (クライアント) に割振られた1つまたはそれ以上の STS-1 のグルーピングにより規定された仮想ポートを利用したインテリジェントパケット転送に関する。

【解決手段】 ソースポートにおける1つの仮想ポートから宛先ポートにおける別の仮想ポートまで、クライアントデータを搬送するためのデータフローパスが確立される。該データフローパスは、データフローパスについて各ポートに通知し、ソースポートに対し宛先ポートを識別するアドレスを提供し、宛先ポートに対しソースポートを識別するアドレスを提供することによって、確立される。



## 【特許請求の範囲】

【請求項1】  $N \times STS-1$  SONEtネットワーク上でソースポートから宛先ポートまで、前記ソースポートにおいて1又はそれ以上の $STS-1$ グループを占有するクライアントデータをパケット転送するための方法において、前記ソースポートから、前記 $STS-1$ により規定された前記宛先ポートまで、前記 $STS-1$ により規定されたデータフローパスを確立する段階を含んだ方法。

【請求項2】 前記データフローパスを各前記ポートに通知する段階と、前記宛先ポートを識別するアドレスを前記ソースポートに提供する段階と、前記ソースポートを識別するアドレスを前記宛先ポートに提供する段階と、を含んだ請求項1に記載の方法。

【請求項3】  $N \times STS-1$  SONEtネットワーク上でソースポートから宛先ポートまで、前記ソースポートにおいて1またはそれ以上の $STS-1$ のグループを占有するクライアントデータをパケット転送するための装置において、

(a) 前記 $STS-1$ からなり、かつ、前記ソースポートから前記宛先ポートに至るデータフローパスを識別し、前記データフローパスを各前記ポートに通知するためのパス識別部と、

(b) 前記宛先ポートを識別するアドレスを識別し、それを前記ソースポートに提供し、かつ、前記ソースポートを識別するアドレスを識別し、それを前記宛先ポートに提供するためのアドレス識別部と、を有する装置。

## 【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、ネットワークアクセスデバイスの物理ポートを仮想ポートに細分するための方法および装置に関する。

【0002】

【従来の技術および発明が解決しようとする課題】 ブリッジーサネット（登録商標）ネットワークは、時として、物理ポートの数によって可能になる以上に多くのブリッジ要素でのピアリング（peer ing）を可能にすべく、物理ポートを仮想ポートに細分することを必要とする。現在、かかる仮想ポートを確立するため通信業界ではATM VPI/VCIおよびフレームリレーパスが用いられている。しかしながら、非常に高い帯域については、これらの技術は技術的理由およびそのコストの高さの両方に起因して適切なものではない。一方、SONET/SDHデバイスは、非常に高い帯域に容易に拡大でき、メトロポリタンエリアネットワークおよび広域ネットワーク（MAN/WAN）において広く用いられている。

【0003】 従って、例えばファイバチャネル（FC）またはギガビットイーサネット（GbE）インタフェー

スから、SONETメトロポリタンエリアネットワーク（MAN）/広域ネットワーク（WAN）まで、といったような、光ラインデータを搬送するネットワークアクセスデバイスを、経済的かつ有効にブリッジするための手段に対するニーズが存在している。

【0004】

【課題を解決するための手段】 本発明によると、 $N \times STS-1$  SONEtネットワークに接続されたソースおよび宛先I/Oデバイス（クライアント）に割振られた1またはそれ以上の $STS-1$ のグルーピングによって規定された仮想ポートを利用するインテリジェントパケット転送のための方法および装置が提供される。クライアントデータを搬送するためのデータフローパスが、ソースポートにおける1つの仮想ポートから宛先ポートにおける他の仮想ポートまで、確立される。該データフローパスは、各ポートに対してデータフローパスについて通知し、ソースポートに対して宛先ポートを識別するアドレスを提供し、宛先ポートに対してソースポートを識別するアドレスを提供することによって、確立される。

【0005】

【発明の実施の形態】 本発明についてここで、全体にわたり同じ参照番号が同じ要素を表わす以下の図面を参照しながら詳細に説明する。

【0006】 図1は、データ搬送ネットワーク150であって、このネットワークを通してあるI/Oデバイス100から出力された光シリアルデータを宛先のI/Oデバイス100まで搬送するためのデータ搬送ネットワーク150、の一例を概略的に示すブロック図である。光高速ポインタツーポインタリンク110を通して、ネットワークアクセスデバイス（NAD）200は、各I/Oデバイス100をネットワーク150に對しインタフェースし、そしてデータフローを制御し、かつ、NADにインタフェースする入力I/Oデバイス100の各々に對し $STS-1$ （すなわち帯域）をフレキシブルに割振るような形で、光シリアルフレームデータ信号を光SONETフレーム信号に多重化する。データ搬送の受信端では、もう1つのNAD200がメトロエリア/広域リンク150から搬送されたSONETフォーマットを受信し、データを分離（デマルチプレクス）し、それを他のI/Oデバイス100への搬送のためもう1つの光高速ポインタツーポインタリンク110上へ出力する。光シリアルリンク110、110aは、イーサネット（例えば10Mbps、100Mbpsまたは10Gbps）またはファイバチャネル（FC）であって良く、メトロエリア/広域ファイバリンク120は、クリアチャネルを提供するため、専用ファイバ、ダークファイバまたはWDM（波長分割多重化）ファイバのいずれであっても良い。

【0007】 図2はデータ搬送ネットワーク150のエン

ンドツーエンドリンクの機能ブロック図、すなわち、NAD200の多重／分離（マクチュプレクサ／デマルチプレクサ）機能を提供するマッパー／アグリゲータ（mapper/aggregator）10およびデマッパー／デアグリゲータ80を含むネットワークの送信（Tx）ノードから受信（Rx）ノードまで、データを搬送するための機能ブロックを示す。この図においては、クライアント（すなわちI/Oデバイス）20i, 20jにサービス提供する2つのNADにより提供される送信（Tx）および受信（Rx）ノードのネットワークコンポーネントが示されている。

【0008】NAD200のマクチュプレクサ／デマルチプレクサコンポーネントについては、本願と共に提出され本願と同じ譲受人が所有する出願の中でより詳しく記述されており、参考として該出願を本願発明の全体に亘って取り入れている。送信ノードにおいては、マッパー／アグリゲータ10は、データを搬送すべく、互いに組合わさってフレキシブル時分割多重（SONET標準）に基づいて、を行う2つの関連した機能を果たす。まず第1にマッパーは、クライアントデータをSONETペイロードにマッピングし、かくして、特定のクライアント20iに提供された帯域割振り構成に基づいて、STS-1が該特定のクライアント20iに割振れることとなる。各STS-1は、固定量の帯域を表わすことから、各クライアントに割当てられたSTS-1の数は、どれほどの帯域がそのクライアントに割振られるかを決定する。有利なことに、各クライアントに対する帯域の割当ては、それが、ユーザ、ネットワークオペレータ、アプリケーションおよび／またはトラヒック、時刻などといったようなネットワーク条件によって設定される任意の数の決定因子（制御パラメータ）に基づいて割当ておよび変更されうという点で、フレキシブルである。第2に、アグリゲータが、クライアントのSTS-1を複合STSペイロード（例えばSTS-48またはSTS-192）に集合させる（すなわち組合せる）。アグリゲータ（集合）データストリームは次に、標準のSONET送信フレーム30に供給されここでセクション／ライン／パスオーバーパッド情報（データ）を挿入して、適切なSONETフレームを形成する。フレーム30からのパラレル出力は次にシリアル化され、モトローラ／広域リンク150の光チャネル50での伝送のためにシリアルライザ／E-Oコンバータ40により、光信号に変換される。

【0009】受信ノードでは受信された光信号が、O-E／デシリアルライザコンバータ60により電気信号ならびにパラレルフォーマットに変換し戻され、その後SONET受信フレーム70内に供給され、ここでセクション／ライン／パスオーバーパッドデータが抽出され処理される。結果として得られたデータストリームは、マッパー／アグリゲータ10の逆機能を果たすデアグリゲータ

／デマッパー80に送出される。デアグリゲータは、SONETフレームから複合STS-48ペイロードを受信し、該デアグリゲータに入力された割当てローカル帯域割振り構成に基づいて、次に、STS-1が割当てられているクライアント20jに従い、複合STSストリームを該STS-1に分離する。このとき、デマッパーは特定のクライアント20jに割当てられたSTS-1を取り出し、これらのSTS-1からクライアントデータを抽出する。

【0010】マッパー／アグリゲータ10およびデアグリゲータ／デマッパー80は、それぞれ、各STS-1を特定のクライアントに割当て／マッピングするために、構成メモリ（configuration memory）を使用する。いずれかの特定の帯域割当て（STS-1）について、NをチャネルのSTS-1の合計数であるとしたとき（例えば、OC-48チャネルが使用される場合には、48）、 $n=0$ から $n=N$ までの任意の数（n）のSTS-1がクライアントに割当てることができる。ネットワーク管理ソフトウェア（ネットワークコントローラ）が、特定のクライアント20i, 20jに対しSTS-1を割当てするための帯域割振りマップを、送信（Tx）および受信（Rx）ノードに提供することができる。

【0011】本発明によると、NADの物理ポートのSTS-1のグルーピングをベースにして作成された2以上の仮想ポート間でのデータトラヒックを方向付け（direct）するために、レイヤ2またはレイヤ3（またはレイヤ1より上の任意のレイヤ）からのフィールドが使用される。図1（b）によって例示されているように、NAD200の物理ポートは、特定のクライアントに割振られたSTS-1のグループからなる仮想ポートへと細分される。STS-1の各グループは、別々の（全く異なる）トラヒックストリームとして処理され、ネットワーク150上で各STS-1の独自の宛先へ独立にルーティングされる。図1（b）に示す例においては、ネットワークアクセスデバイスNAD1は3つの物理ポート140を有し、第1のポート145は2つの仮想ポート1Aと1Bからなり、各該ポートは、データがこれらのSTS-1グループによって搬送される特定のI/Oデバイスに割振られるある数のSTS-1からなる。与えられたSTS-1グループのためのフローパス、（S）の各エンドにおけるレイヤアドレスリングについての知識（例えばイーサネットリンク間のレイヤ2については、これはMACアドレスリングとなる）が、これらのエンドポイント間のブリッジ（ここでは「仮想ブリッジ」と呼ばれる）を規定するために、使用される。有利には、この機能性は集成的に、ネットワーク全体にわたるフレキシブルブリッジを提供する。

【0012】図3の表は、ネットワークアクセスデバイスNAD2およびNAD3の各々に対し、ネットワーク

アクセスデバイスNAD1の第1の物理ポート145によって出力されるデータを転送するため、ネットワーク内において適用される一組の規則を例示している。この例については、イーサネットデータが仮定されており、Sバスの終端(エンド)のアドレッシングのために使用されるフィールドは、レイヤ2MACアドレスである(代替的にはIPアドレスを使用することもできる)。STS-1が割り振られたクライアントのグループによって仮想ポート1Aが規定され、その宛先はNAD2から供給を受けるI/Oデバイス100である。またもう1つの仮想ポート1Bは、STS-1が割り振られたクライアントの異なるグループによって規定され、その宛先はNAD3から供給を受けるもう1つのI/Oデバイス100である。図1および3に示す例では、NAD2にインタフェースされた宛先I/Oデバイスは、ルータR2であり、NAD3にインタフェースされた宛先I/Oデバイスは、もう1つのルータR3である。一例として、図3では、R2のためのMACアドレス(レイヤ2)はM2と称し、R3のためのMACアドレスはM3と称す(1P2と称されるR2のためのIPアドレスおよび1P3と称されるR3のためのIPアドレスもまた、レイヤ3アドレスを代わりに使用できることを示すために、例示されている)。

【0013】この例については、仮想ポート1A~R2によって規定されるデータおよび仮想ポート1B~R3によって規定されるデータを転送するためにネットワーク内にはある規則セットが適用されており、これらは図3の表(c)によって示されている。例示するとおり、各仮想ポートのフローを方向付けるために、宛先I/Oデバイスに割当てられたMACアドレスが使用される。

【0014】一例を用いて本発明のルーティング方法を示すため、以下のような選択された定義が用いられる。

【0015】(a) 単一のパイプとして扱われコアネットワークを通して一緒にルーティングされるSTS-1のグループを、s-バスと称する。

【0016】(b) ネットワーク上の3つのI/Oデバイス(1/O1、1/O2および1/O3)のMACアドレスを、MA1、MA2およびMA3と称する。

【0017】(c) 3つのI/Oデバイスは3つの異なるNAD上にある。1/O1はNAD1上にあり、1/O2はNAD2上にあり、1/O3はNAD3上にある。

【0018】(d) ルーティングされるべき異なるs-バスが3つ存在し、各s-バスは2つのI/Oデバイス間のトラフィックフローを表わす。S1-2は、1/O1と1/O2の間のトラフィックフローを表わし、このフローを搬送する。S1-3は、1/O1と1/O3の間のトラフィックフローを表わし、このフローを搬送する。S2-3は、1/O2と1/O3の間のトラフィックフローを表わし、このフローを搬送する。

【0019】ユーザが1/O1と1/O2との間のネット

ワークを通して接続を確立するためには、s-バスS1-2が作成されなくてはならない。この時点でNAD1およびNAD2の双方は、S1-2に加えてs-バスの終端(エンド)におけるMACアドレスも認知していなくてはならない。このs-バス対MACアドレスの割当ては、以下のとおり、ネットワーク管理システム(NMS)によって確立され得る。

【0020】1. NMSは、ユーザから、1/O1が(そのMACアドレスMA1を通して)NAD1に付帯することおよび1/O2がNAD2を通して付帯することを、知らされる。

【0021】2. 1/O1と1/O2との間のs-バスS1-2が作成されたときに、NMSはNAD1に対し、MA2がS1-2の終端にあることを知らせ、NMSはNAD2に対し、MA1がS1-2の終端にあることを知らせる。

【0022】代替的には、MACアドレス対s-バスの割当てを確立するためにNMSを使用する代わりに、ユーザは以下のように直接それを行うことができる。

【0023】1. ユーザは、新しいs-バスS1-2がネットワーク内に存在することをNAD1に対し直接知らせる。

【0024】2. ユーザは、MA2がS1-2の終端にあることをNAD1に知らせる。

【0025】3. ユーザは、新しいs-バスS1-2がネットワーク内に存在することを直接NAD2に知らせる。

【0026】4. ユーザは、MA1がS1-2の終端にあることをNAD2に知らせる。

【0027】多数のI/Oデバイスを単一のNADに接続できるということに留意されたい。上述のs-バスとMACアドレスの割当ての全てにおいて、I/OデバイスおよびそのMACアドレスは、NAD全体ではなく、NAD上のあるポートに関連付けられる。

【0028】NMS/ユーザが、I/OデバイスおよびこのI/Oデバイスに付随するNADに対し、作成されたs-バスおよびその終端にあるMACアドレスを知らせるために、さまざまな手段を使用することができる。この情報は、インバンド(すなわちSONETベロード/オーバーヘッド内)かまたはアウトバンド(すなわち外部ネットワークを介してNADに搬送される)かのいずれかで提供され得る。

【0029】ネットワーク150は、分散I2スイッチを提供する。STS-1フローにおける部分メッシュのみに可能性があることから、あるI/Oデバイスポートのみが接続可能である(例えば、もしSTS-1フローが第1および第2のI/Oデバイスポートの間に存在しないならば、これらのポートは接続され得ない)。このような場合、第1のI/Oデバイスポートから第2のI/Oデバイスポートに至るデータトラフィックは、第2の

I/OデバイスポートまでのSTS-1フローをもたない他のI/Oポートを介してルーティングされなくてはならず、かかる延長ルートの計算は、サポートされたルーティングプロトコルによって自動的に行われる。また、MACアドレス対NADポートの割当ては手動で行われることから、この方法は、本質的にセキュアであり、そのため、第三者がsーパスグループを通して作成される私設ネットワークに「侵入することまたはそれを「欺く」ことを困難にしている。

【0030】有利には、イーサネットフレームのあるフローに対しSTS-1のグループを割当てる上述の方法は、QoSが結果として自動的にもたらされることを保証する。そのため、帯域についてのいかなるコンテンションもなくコアが実現されるような形で、非QoSウェア (aware) イーサネット/I PSスイッチによるトラフィック保証の形態を、ITネットワークが実現することができる。

【0031】上述の好ましい実施形態で利用される個々の電子および処理機能は、個々に、当業者には充分理解されているものである。さまざまなその他の実現が、代替として当業者により考えることを理解すべきである。通信設計の分野の当業者であれば、本発明のある利用分野のための適切な実現方法に容易に応用できるであろう。

【0032】従って、本明細書に例として示され記述された特定の実施形態が、特許請求の範囲によって規定されるクレームされた発明の範囲を制限すべく意図されたものでないということを理解すべきである。

#### 【図面の簡単な説明】

【図1】(a)は、ネットワークアクセスデバイス(NAD)内に備わる本発明の装置および方法の機能性を仮想バスにより例示するメトロエリア/広域データ搬送\*

\* ネットワークの概略ブロック図であり、(b)は、ネットワークアクセスデバイスNAD1の物理ポートによって提供される仮想ポートを概略的に示す図である。

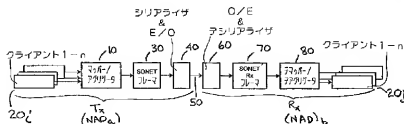
【図2】ネットワークの送信(Tx)ノードからデータを搬送するためのネットワークアクセスデバイスの送信側コンポーネントと、ネットワークの受信(Rx)ノードにおいてデータを受信するための他のネットワークアクセスデバイスの受信側コンポーネントとを示す、図1のデータ移送ネットワークのエンドツーエンドリンクの機能ブロック図である。

【図3】(a)、(b)、および(c)は、全て、NAD1がイーサネットリンクからデータを送出する例を特定の状況についてのものであり、I/Oデバイスを識別する1組のMAC/IPアドレス(a)、特定の宛先I/Oデバイスのための1組のアドレス(b)および図1に示すネットワークアクセスデバイスNAD1により出力されたデータの転送のための1組の規則(IPアドレスではなくむしろMACアドレスが選択された規則)(c)を例示する表である。

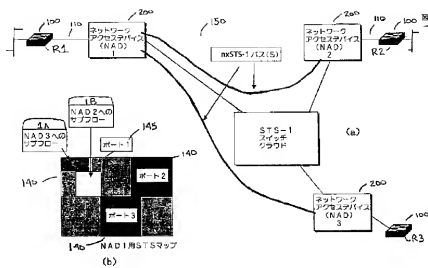
#### 【符号の説明】

- 10…マッパー/アグリゲータ
- 30…SONET送信フレーム
- 40…シリアルライザ/E-Oコンバータ
- 50…光チャネル
- 60…O-E/デシリアルライザコンバータ
- 70…SONET受信フレーム
- 80…デアグリゲータ/デマッパー
- 100…I/Oデバイス
- 140…物理ポート
- 150…データ搬送ネットワーク
- 200…ネットワークアクセスデバイス(NAD)

【図2】



【図1】



【図3】

図 3

ルータ名	MAC アドレス	IP アドレス
R1	M1	IP1
R2	M2	IP2
R3	M3	IP3

(a)

宛先ルータ	IPアドレス (レイヤ3)	MACアドレス (レイヤ2)
R2	IP2	M2
R3	IP3	M3

(b)

宛先MAC	転送決定
同接連	全サブフロー
M2	2へのSTS-1サブフロー
M3	3へのSTS-1サブフロー
未知	ドロップフレーム

(c)

フロントページの続き

(71)出願人 501459712

150 Katimavik Road, S  
uite 2000 Kanata, Onta  
rio K2L 2N2 CANADA

Fターム(参考) 5K030 GA08 HA01 HB13 HB14 HD09  
JA01 JL03 JL10 KA05 KA13  
LB01 LC01 LC06

## 【外国語明細書】

## 1. Title of the Invention

## METHOD AND APPARATUS FOR PACKET FORWARDING OVER NXSTS-1 SONET

## 2. Detailed Description of the Invention

This invention relates to a method and apparatus for subdividing the physical ports of network access devices into virtual ports.

A bridged ethernet network sometimes requires that a physical port be subdivided into virtual ports to allow peering with more bridging elements than the number of physical ports allows. Currently, ATM VPs/VCs and Frame Relay paths are used in the communications industry to establish such virtual ports. However, for very high bandwidths these technologies are not suitable due to both technical reasons and their high cost. On the other hand SONET/SDH devices can easily scale to very high bandwidths and are widespread in metropolitan and wide area networks (MAN/WAN).

There is a need, therefore, for means to economically and effectively bridge network access devices transporting optical line data, such as from a fiber channel (FC) or Gigabit Ethernet (GbE) interface, to a SONET metropolitan area network (MAN) wide area network (WAN).

In accordance with the present invention there is provided a method and apparatus for intelligent packet forwarding utilizing virtual ports defined by a grouping of one or more STS-1s allocated to source and destination I/O devices (client) connected to an N x STS-1 SONET network. A data flow path for transporting the client data is established from one virtual port at the source port to another virtual port at the destination port. The data flow path is established by notifying each of the ports of the data flow path and providing to the source port



an address identifying the destination port and providing to the destination port an address identifying the source port.

The invention will now be described in detail with reference to the following drawings in which like reference numerals refer to like elements throughout.

### 3. Brief Description of the Drawings

Figure 1(a) is a schematic block diagram of a metro/wide area data transport network illustrating, by the virtual paths S, the functionality of the method and apparatus of the present invention which is comprised within network access devices (NADs) and Figure 1(b) is a schematic illustration of the virtual ports provided by a physical port 1 of the network access device NAD1;

Figure 2 is a functional block diagram of an end-to-end link of a the data transport network of Figure 1 showing the components of the transmit side of a network access device, for transporting data from a transmit (Tx) node of the network, and the components of the receive side of another network access device, for receiving data at a receive (Rx) node of the network; and,

Figures 3 (a), (b) and (c) are tables which illustrate a set of MAC/IP addresses identifying I/O devices (a), a set of addresses for the particular destination I/O devices of the example (b) and a set of exemplary rules (for which the MAC address is chosen for use rather than the IP address) for the forwarding of data output by network access device NAD1 shown in Figure 1 (c), all for the particular circumstances of the example chosen in which NAD1 is sending data from an ethernet link.

Figure 1 of the drawings is a schematic block diagram illustrating an exemplary data transport network 150 for transporting optical serial data output

from one I/O device 100 for transport through the network 150 to another I/O device 100. Through a high speed optical point-to-point link 110 a network access device (NAD) 200 interfaces each I/O device 100 to the network 150 and multiplexes the optical serial frame data signal into an optical SONET frame signal in a manner which controls the flow of the data and flexibly allocates STS-1s (i.e. bandwidth) to each of the input I/O devices 100 interfaced to the NAD. At the receive end of the data transport another NAD 200 receives the transported SONET format data from the metro/wide area link 150 and demultiplexes and outputs the data onto another high speed optical point-to-point link 110 for transport to the other I/O device 100. The optical serial links 110, 110a could be ethernet (e.g. 10 Mbps, 100 Mbps or 10Gbps) or Fiber Channel (FC) and the metro/wide area fiber link 120 could be any of a dedicated fiber, dark fiber or WDM (Wavelength Division Multiplexed) fiber to provide a clear channel.

Figure 2 is a functional block diagram of an end-to-end link of the data transport network 150 showing the functional blocks thereof for transporting data from a transmit (Tx) node to a receive (Rx) node of the network including the mapper/aggregator 10 and demapper/deaggregator 80 which provide the multiplexer/demultiplexer functions of the NAD 200. In this figure the network components of the transmit (Tx) and receive (Rx) nodes provided by two NADs, servicing clients (i.e. I/O devices) 20<sub>1</sub>, are shown.

The multiplexer/demultiplexer components of the NAD 200 are described in greater detail in an application filed herewith and owned by same assignee as this application, which is incorporated herein by reference. At the transmit node the mapper/aggregator 10 performs two related functions which combine to perform flexible time-division multiplexing (based on the SONET standard) to transport the data. First, a mapper maps the client data into a SONET payload whereby STS-1's are allocated to a particular client 20, based on a bandwidth allocation configuration provided to it. Since each STS-1 represents a fixed amount of bandwidth the

number of STS-1's assigned to a client determines how much bandwidth is allocated to that client. Advantageously, the assignment of bandwidth to each client is flexible in that it can be assigned and changed based on any number of determinants (control parameters) set by a user, a network operator, an application and/or network conditions such as traffic, time of day, etc. Second, an aggregator aggregates (i.e. combines) the client STS-1's into a composite STS payload (e.g. STS-48 or STS-192). The aggregate data stream is then fed into a standard SONET transmit framer 30 to insert section/line/path overhead information (data) and create a proper SONET frame. The parallel output from the framer 30 is then serialized and converted to an optical signal by a serializer/E-O converter 40 for transmission over an optical channel 50 of the metro/wide area link 150.

At the receive node the received optical signal is converted back to an electrical signal and to parallel format by an O-E/deserializer converter 60 and then fed into a SONET receive framer 70 where the section/line/path overhead data is extracted and processed. The resulting data stream is passed into a de-aggregator/de-mapper 80 which performs the inverse function of the mapper/aggregator 10. The deaggregator receives the composite STS-48 payload from the SONET framer and, based upon an assigned local bandwidth allocation configuration input thereto, it then separates the composite STS stream into STS-1's according to the client 20, they are assigned to. The demapper then takes the STS-1's assigned to a particular client 20, and extracts the client data from them.

The mapper/aggregator 10 and de-aggregator/de-mapper 80 each use a configuration memory to assign/map each STS-1 to a particular client. For any particular assignment of bandwidth (STS-1s) a client can be assigned any number (n) of STS-1s from  $n=0$  to  $n=N$ , where N is the total number of STS-1s of the channel (e.g. 48 if an OC-48 channel is used). Network management software (a network controller) may provides the transmit (Tx) and receive (Rx) nodes with the

bandwidth allocation map which assigns STS-1's to the particular clients  $20_i$ ,  $20_j$ .

In accordance with the present invention fields from layer 2 or layer 3 (or any layer above layer 1) are used to direct data traffic between virtual ports created on the basis of groupings of STS-1s of the physical ports of the NAD. As illustrated by Figure 1(b) the physical ports of an NAD 200 are subdivided into virtual ports made up of the groups of STS-1s allocated to a particular client. Each group of STS-1s are treated as a separate (distinct) traffic stream and are independently routed over the network 150 to their own destination. In the example shown by Figure 1(b) the network access device NAD1 has three physical ports 140 and the first port 145 comprises two virtual ports 1A and 1B each comprising the number of STS-1s which are allocated to the particular I/O devices whose data is carried by those STS-1 groups. Knowledge of the layer addressing (e.g. for layer 2 in an ethernet link this would be MAC addressing) at each end of the flow path (S) for a given STS-1 group is used to define a bridge between those end points (referred to herein as a "virtual bridge"). Advantageously, this functionality collectively provides flexible bridging over the entire network.

The tables of Figure 3 illustrate a set of exemplary rules applied within the network for the forwarding of data output by a first physical port 145 of the network access device NAD1 to each of network access devices NAD2 and NAD3. For this example ethernet data is assumed and the field used for the addressing of the ends of S-paths is the layer 2 MAC address (although the IP address could alternatively be used). A virtual port 1A is defined by a group of client allocated STS-1s whose destination is an I/O device 100 fed by NAD2 and another virtual port 1B is defined by a different group of client allocated STS-1s whose destination is another I/O device 100 fed by NAD3. In the example shown in Figures 1 and 3 the destination I/O device interfaced to NAD2 is a router R2 and the destination I/O device interfaced to NAD3 is another router R3. For illustration purposes in Figure 3 the MAC address (layer 2) for R2 is designated M2 and the MAC address for R3

is designated M3. (The IP address for R2 designated as IP2 and the IP address for R3 designated as IP3 are also illustrated to indicate that layer 3 addresses can instead be used.)

For this example a given set of rules is applied within the network in order to transport the data defined by virtual port 1A to R2 and the data defined by virtual port 1B to R3 and these are shown by the table (c) of Figure 3. As illustrated the MAC address assigned to the destination I/O device is used to direct the flow of each virtual port.

For purposes of illustrating the routing method of the invention, using an example, the following selected definitions are used:

- (a) The group of STS-1s that are treated as a single pipe and routed through the core network together are referred to as an s-path.
- (b) The MAC addresses of three I/O devices (I/O1, I/O2 and I/O3) on the network are referenced MA1, MA2 and MA3.
- (c) The three I/O devices are on three different NADs. I/O1 is on NAD1, I/O2 is on NAD2 and I/O3 is on NAD3.
- (d) There are three different s-path groups to be routed and each s-path represents a traffic flow between two I/O devices. S1-2 represents and carries the traffic flow between the I/O1 and I/O2. S1-3 represents and carries the traffic flow between I/O1 and I/O3. S2-3 represents and carries the traffic flow between I/O2 and I/O3.

For a user to establish a connection through the network between I/O1 and I/O2, the s-path S1-2 must be created. At this point, both NAD1 and NAD2 must be made aware of S1-2 and also the MAC addresses at the ends of the s-path. This s-path to MAC addressment assignment may be established by the network management system (NMS) as follows:

1. The NMS is informed by the user that I/O1 (through it's MAC address MA1) is attached to NAD1 and that I/O2 is attached through NAD2.
2. When the s-path S1-2 between I/O1 and I/O2 is created, the NMS informs NAD1 that MA2 is at the end of S1-2 and the NMS informs NAD2 that MA1 is at the end of S1-2.

Alternatively, instead of using the NMS to establish the MAC to s-path assignment, a user can do so directly as follows:

1. The user informs NAD1 directly that a new s-path S1-2 exists in the network.
2. The user informs NAD1 that MA2 is at the end of S1-2.
3. The user informs NAD2 directly that a new s-path S1-2 exists in the network.
4. The user informs NAD2 that MA1 is at the end of S1-2.

It is to be noted that multiple I/O devices may be connected to a single NAD. In all of the above s-path and MAC address assignments, the I/O device and it's MAC address are associated with a port on the NAD and not the entire NAD.

Various means may be used by the NMS/user to inform the I/O devices and NADs associated therewith of the created S-paths and MAC addresses at the ends thereof. This information can be provided either in-band (i.e. within the SONEt payload/overhead) or out-of-band (i.e. carried to the NADs via an external network).

The network 150 provides an appearance of a distributed L2 switch. As there is a potential for only a partial mesh of STS-1 flows only certain I/O device ports will be connectable (e.g. if an STS-1 flow does not exist between a first and second I/O device port those ports cannot be connected). In such case the data traffic going from a first I/O device port to a second I/O device port must be routed via another I/O port which does have an STS-1 flow to the second I/O device port and calculation of such extended route is accomplished automatically by supported

routing protocols. Also, as the MAC address to NAD port assignment is done manually, this method is inherently secure, making it difficult for third-parties to "break into" or "spoof" a private network created through a group of s-paths.

Advantageously, the foregoing method of assigning a group of STS-1s to a given flow of ethernet frames guarantees that CoS automatically results. As such the IT network is able to implement a form of traffic guarantee with non-QoS-aware ethernet/IP switches such that a core is implemented with no contention for bandwidth.

The individual electronic and processing functions utilised in the foregoing described preferred embodiment are, individually, well understood by those skilled in the art. It is to be understood by the reader that a variety of other implementations may be devised by skilled persons for substitution. Persons skilled in the field of communication design will be readily able to apply the present invention to an appropriate implementation method for a given application.

Consequently, it is to be understood that the particular embodiment shown and described herein by way of illustration is not intended to limit the scope of the invention claimed by the inventors which is defined by the appended claims.

#### 4. Claims

1. A method for packet forwarding over an N x STS-1 SONET network from a source port to a destination port client data occupying a group of one or more STS-1s at said source port, said method comprising establishing a data flow path defined by said STS-1s from said source port to said destination port defined by said STS-1s.
2. A method according to claim 1 comprising notifying each of said ports of said data flow path, providing to said source port an address identifying said destination port and providing to said destination port an address identifying said source port.
3. Apparatus for packet forwarding over an N x STS-1 SONET network from a source port to a destination port client data occupying a group of one or more STS-1s at said source port, said apparatus comprising:
  - (a) a path identifier for identifying a data flow path comprising said STS-1s and extending from said source port to said destination port and notifying each of said ports of said data flow path; and,
  - (b) an address identifier for identifying and providing to said source port an address identifying said destination port and for identifying and providing to said destination port an address identifying said source port.



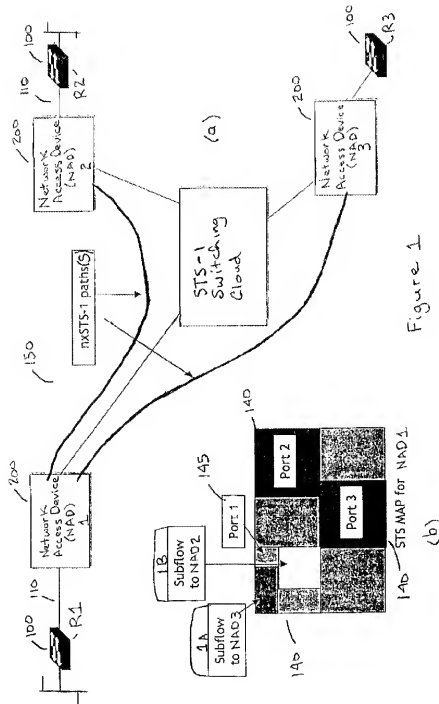


Figure 1

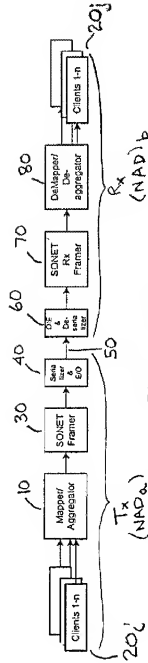


Figure 3

Router Name	MAC Address	IP Address
R1	M1	IP1
R2	M2	IP2
R3	M3	IP3

(a)

Router 1's Route and ARP table

Destination Router	IP Address (Layer 3)	MAC Address (Layer 2)
R2	IP2	M2
R3	IP3	M3

(b)

NAD1's forwarding rule table (static)

Destination MAC	Forwarding Decision
Broadcast	All sub-flows
M2	STS-1 subflow to 2
M3	STS-1 subflow to 3
Unknown	Drop frame

(c)

## 1. Abstract

A method and apparatus for intelligent packet forwarding utilizing virtual ports defined by a grouping of one or more STS-1s allocated to source and destination I/O devices (client) connected to an N x STS-1 SONET network. A data flow path for transporting the client data is established from one virtual port at the source port to another virtual port at the destination port. The data flow path is established by notifying each of the ports of the data flow path and providing to the source port an address identifying the destination port and providing to the destination port an address identifying the source port.

## 2. Representative Drawing

Figure 1